

특 2000-0047589

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(51) Int. Cl.⁷
G06F 3/14

(11) 공개번호 특2000-0047589
(43) 공개일자 2000년07월25일

(21) 출원번호	10-1999-0048295
(22) 출원일자	1999년11월03일
(30) 우선권주장	09/213,857 1998년12월16일 미국 (US)
(71) 출원인	인터내셔널 비지네스 머신즈 코포레이션 포만 제프리 엘
	미국 10504 뉴욕주 아몽크
(72) 발명자	탄넨바움알렌리차드
	미국78731텍사스주오스틴그레이스톤드라이브3801
(74) 대리인	김창세, 김원준, 장성구

심사청구 : 있음

(54) 음성 커맨드 인식 컴퓨터 시스템에서 디스플레이 상에 피드백을 제시하는 방법 및 장치

요약

본 발명은 음성 가동 컴퓨터 시스템에 관한 것으로서, 디스플레이 스크린에 상의 이산 위치(discrete location)와 관련되거나 그 위치로부터 프롬프트되는 단일 사용자(an end user)로부터의 음성 커맨드가 이 시스템에 의해 분석된다. 이에 응답하여, 음성 입력의 분석된 내용 및 문맥과 기능적으로 관련된 위치 또는 그 발음을 프롬프트하는 스크린 위치에서 확인 영역이 스크린 상에 디스플레이된다. 확인 영역 내에는 그 발음의 컴퓨터 해석이 선택적으로 조절가능한 속도와 시간만큼 지속 및 소멸 디스플레이된다. 그에 따라, 인식된 발음은 사용자 초점이 존재할 것으로 여겨지는 가변 위치에서 확인 영역내에 디스플레이된다. 디스플레이 스크린상의 다른 내용을 가리거나 및/또는 디스플레이 스크린 상의 현재 관심 위치로부터 확인 영역이 디스플레이되는 다른 위치로 시선을 이동시키도록 요구함으로써 단일 사용자 초점을 훼손하는 고정 위치 확인 영역과 연관된 산만함(distractions)을 피할 수 있다. 확인 영역과 그 내부에 디스플레이된 연관된 해석 음성 커맨드의 지속 및 소멸 시간은 그 커맨드가 정확하게 분석된 음성 인식 시스템의 신뢰도에 따라 자동으로 변할 수 있고, 올바르게 인식된 커맨드는 확인 윈도우 내에 디스플레이된 후 자동으로 실행된다.

도면

도3

참제서

도면의 간단한 설명

도 1은 확인 영역을 디스플레이하는 종래 기술의 방법을 나타내는 음성 가동 컴퓨터 시스템에서의 대표적인 디스플레이 스크린,

도 2a 및 도 2b는 음성 커맨드 및 제어 시스템과 연관된 확인 영역의 개선된 디스플레이를 위한 본 발명에 따른 디스플레이 스크린을 나타내는 도면,

도 3은 도 2에 도시한 개선된 확인 영역을 실행하기 위해 본 발명의 컴퓨터 시스템에서 실행되는 프로그램 코드에서 구현가능한 흐름도,

도 4는 본 발명의 특징이 유리하게 구현되는 음성 가동 컴퓨터 시스템을 묘사하는 단순화된 블록도.

도면의 주요 부분에 대한 부호의 설명

- 10 : CPU. 11 : ROM
- 12 : RAM 13 : I/O 어댑터
- 14 : 통신 어댑터 15 : DASD
- 16 : 공통 버스 17 : 키보드
- 18 : 사용자 인터페이스 어댑터 19 : 디스플레이 어댑터

- 21 : 디스플레이 23 : 조이 스틱
25 : 마우스 27 : 마이크로폰

발명의 상세한 설명

발명의 목적

발명이 속하는 기술분야 및 그 분야의 종래기술

본 발명은 음성 커맨드 컴퓨터 시스템에 관한 것으로서, 보다 구체적으로는 단일 사용자에게 커맨드 피드백을 제시하는 시스템에 관한 것이다.

컴퓨터 분야가 성숙함에 따라, 단일 사용자가 시스템과 인터페이스하는 방식은 상당히 개선되었다. 원래는 단일 사용자의 키보드 입력을 요구하는 커맨드 라인 인터페이스였던 것이, 이제는 컴퓨터와의 보다 자연스러운 휴먼 인터페이스를 촉진하는 다양한 기법과 장치가 존재한다. 그 대표적인 예로는 마우스, 트랙볼, 터치스크린과 같은 다양한 지시 장치(pointing devices)와 컴퓨터 음성 인식의 진보로 인해 가능하게 된 음성 자동 인터페이스(voice activated interfaces)가 있다.

이러한 음성 인식은 두 가지 주된 범주로 나뉜다. 먼저, 음성 데이터 입력을 위해 개발된 시스템을 들 수 있는데, 그 예로는 IBM 코퍼레이션(IBM Corporation)과 드래곤 시스템(Dragon Systems)에 의해 시장화된 것들과 같은 음성 받아쓰기 또는 음성을 문장으로 바꾸는 시스템(speech dictation or speech-to-text systems)이 있다. 컴퓨터화된 음성 인식 기법의 두 번째 범주는 단일 사용자의 음성 패턴이 검출되고 분석되는 커맨드 및 제어 시스템에 있다. 이러한 커맨드 및 제어 기법의 특수한 형태로는 데스크탑 네비게이터(a desktop navigator)가 있는데, 여기서 사용자는 메뉴 커맨드 단어, 버튼상의 라벨(labels on buttons), 마우스 커서를 상하좌우로 이동시키는 것과 같은 방향 커맨드(directional commands) 등을 단순히 구술함으로써 컴퓨터 시스템을 통해 그래픽 사용자 인터페이스(graphical user interface : GUI)를 네비게이션(navigate)할 수 있다.

이러한 음성 인식 시스템 교육의 음성 인식 오류로 인해 - 부분적으로는 기술 수준으로 인해, 부분적으로는 사용자 구술 패턴의 다양성과 발음되는 올바른 입력의 메모리로 인해 -, 이들 시스템은 전형적으로 사용자 발음의 시스템 해석을 디스플레이하기 위한 전용의 사전결정된 GUI 영역을 보유한다. 사용자는 이 영역을 조사함으로써 자신의 발음이 실제로 올바르게 해석되었는지 또는 반영되었는지를 확인한다.

현재의 음성 인식 시스템은 GUI 스크린 상의 고정된 위치에 있는 제어 윈도우내에 이 최선의 추측(best guess)을 디스플레이한다. 이 제어 윈도우의 공간적인 배치는 구현에 따라 단일 사용자가 원하는 대로 변경될 수도 있고, 또는 실행 중인 현재 애플리케이션의 타이틀 바(title bar)와 같이 고정된 위치에 배치될 수 있지만, 그럼에도 불구하고 이 기법과 관련된 많은 문제점이 남아 있다. 이를 해석된 커맨드를 디스플레이하는 방법은 그 애플리케이션의 다른 중요 부분의 시계(visibility)를 가리지 않도록 하는 위치에 확인 영역을 배치하려 시도하고 있다. 그러나, 그렇게 함으로써 사용자가 초점을 맞추고 있는 음성 인식의 확인 영역인 스크린의 영역으로부터 시선을 이동시켜야 하는 산만함(distracting)과 혼란(disrupting)을 야기할 수 있는데, 이에 따라 사용자는 확인 영역의 디스플레이에 의해 초래된 중단(interruption) 이전에 사용자의 초점이었던 스크린의 영역을 상기하고 그곳에 자신의 시선을 재배치하여야 한다. 또한, 현재 시스템은 음성 인식 시스템의 결과가 보고되기만 하며 커맨드의 해석을 고려하지 않고 그 결과를 디스플레이하는 문제점이 있다.

이러한 음성 인식 컴퓨터 시스템을 이용하는 중에 이들 단점은 신체적 및 정신적으로 극도로 귀찮은 것일 수 있고, 음성 인식 또는 네비게이션 시스템에 대한 전술한 장점을 거의 무효화하는 방식으로 단일 사용자의 생산성에 심각한 영향을 미칠 수 있다는 것을 쉽게 이해할 것이다. 현재의 방법 및 기법으로는 디스플레이의 중요 영역의 시계를 유지하는 것과 앞서 설명한 초점 및 재초점(refocus)이라는 이들 심각한 문제점을 해결할 수 없다.

따라서, 단일 사용자의 산만성과 GUI 사용자 인터페이스 상에 제시된 애플리케이션의 시계의 열화를 방지하는 음성 커맨드 및 제어 시스템이 고도로 요구된다.

발명이 이루고자 하는 기술적 과제

음성 자동 컴퓨터 시스템에서는, 디스플레이 스크린상의 이산 위치(discrete location)와 관련되거나 그 위치로부터 프롭트되는 커맨드 및/또는 제어 발음과 같은 단일 사용자로부터의 음성 입력이 분석된다. 이러한 분석에 응답하여, 이 음성 입력의 분석된 내용(contents)과 기능적으로 관련된 위치 또는 그 발음이 프롭트된 스크린 위치에서 확인 영역(confirmation area)이 스크린 상에 디스플레이된다. 확인 영역내에는 그 발음의 컴퓨터 해석이 사전선택된 시간 동안 디스플레이되고, 이어서 사전선택된 속도로 점차 소멸되는데, 이러한 지속 및 소멸은 단일 사용자가 선택적으로 조절할 수 있다. 이러한 방식으로, 커맨드나 제어 발음의 해석은 확인 영역 내에 디스플레이되며, 이 확인 영역은 그 커맨드나 제어 발음의 내용과 기능적으로 관련된 가변 위치에서 디스플레이 스크린 상에 공간적으로 배치되어 사용자 시선의 초점이 존재했었고 그 발음을 유발하기도 한 디스플레이 스크린 상의 위치 가까이 있도록 한다. 이렇게, (1) 디스플레이 스크린 상의 관심 내용을 가리거나 및/또는 (2) 디스플레이 스크린 상의 현재 관심 위치로부터 확인 영역이 디스플레이되는 다른 위치로 시선을 이동시키도록 요구함으로써 단일 사용자 초점을 훼손하는 고정 위치 확인 영역과 연관된 산만성이 방지된다. 본 발명에서 중요한 요소는 발음의 해석, 즉 커맨드의 진정한 목적의 해석이다. 피드백은 애플리케이션 기능에 대한 올바른 해석이 결정될 때까지 대부

본의 경우 잠깐 보류된다.

보다 일반적인 경우, 본 발명을 커서 목표(cursor targets), 메뉴 항목(menu items), 지시 장치(pointing devices), 툴바(toolbars)와의 상호작용 및 연관(association)으로 이루어지는 종래 기술의 프로그래머 그래픽 모델(programmers graphical model)에 제한하려는 의도는 아니다. 오히려, 본 발명은 인식된 구술 커맨드를 자연 언어의 형태로 디스플레이하고, 피드백 동작을 그 구술 커맨드를 유발하고 프로토타입하는 그래픽 사용자 인터페이스 내의 초점 영역에 근접하게 하는 시각 메커니즘을 시도하고 있다. 이러한 피드백 동작은 원하는 초점이나 목표 영역에서 올바른 구술 동작(spoken action)이 취하여 졌다는 피드백을 단말 사용자에게 제공하기 위해, 문장 위의 문장(text over text), 시각 정보(visuals), 색채 변화(color change), 애니메이션(animation), 자동으로 목표 영역으로 페이딩(fading)되는 영상이나 메시지의 점진적인 축소를 포함할 수 있다.

발명의 구성 및 작용

먼저 도 1을 참조하면, 종래 기술에 따른 도 4의 시스템의 모니터 상에서 일어날 수 있는 대표적인 디스플레이(30)를 단순화하여 도시하고 있다. 이 시스템에서, 사용자는 도 4의 시스템에서 분석되고 실행될 음성 커맨드를 프롬프트(prompt)할 수도 있는 스크린(30)의 영역(32)과 상호작용하고, 따라서, 여기에 시선의 초점을 맞추고 있을 것이다. 일단 음성 인식 시스템이 이러한 발음을 인식하려고 시도하면, 확인 영역(34) 내에 이러한 커맨드의 컴퓨터 해석이나 번역을 디스플레이하는 것이 통상적이다. 예를 들어, 사용자가 사전 열기(open dictionary)와 같은 음성 가동가능 선택 메뉴(menu of voice activatable choices)가 있는 애플리케이션의 영역(32)을 바라보고 있다면, 사용자는 결국 이 커맨드를 발음할 것이다. 이에 따라, 컴퓨터 시스템이 그 커맨드를 인식한 후 또는 해석을 위한 최선의 추측을 한 후, (발음이 올바르게 해석되었다면 사전 열기와 같은) 이 해석이 이 확인 영역(34) 내에 나타날 것이다.

이전에 언급한 음성 커맨드 시스템의 이러한 동작에 있어서의 중대한 결점은 도 1의 이 종래 기술의 스크린(30)을 참조하면 분명하게 드러난다. 단말 사용자 초점은 영역(32) 상에 있지만, 이 영역에서 프롬프트되어 사용자가 음성 커맨드를 발하였을 때 이 사용자가 (이 커맨드가 어떻게 해석되었는가에 관한) 피드백을 받게 될 장소는, 앞서 커맨드를 프롬프트시킨 영역(32)으로부터 사용자 인터페이스 스크린 상에서 상당한 거리만큼 공간적으로 떨어져 있는 확인 영역(34) 내에서 받게 된다는 것을 알 수 있다. 이로 인해, 사용자는 음성 커맨드가 올바르게 해석되었는지 여부를 식별하기 위해 확인 영역(34)에서 스크린(30)의 대각선 반대 끝으로 자신의 시선의 초점을 다시 맞춰야 한다. 그렇다면, 사용자의 시선은 영역(30)의 대각선 반대 끝으로 자신의 시선의 초점을 다시 맞춰야 한다. 그 결과, 사용자의 시선은 영역(30)의 대각선 반대 끝으로 자신의 시선의 초점을 다시 맞춰야 한다. 통상, 상당시간 동안 지속되는 이 확인 영역(34)을 스크린(30) 상의 고정되고 벗어난 위치에 배치하는 역사적인 한 이유는 애플리케이션의 다른 정보 내용이 디스플레이되는 스크린(30)의 다른 관련 부분을 가리지 않도록 하기 위함이다.

도 2a 및 도 2b를 참조하면, 도 1의 종래 기술 스크린(30)의 예에 대응하는 스크린(30)의 유사한 예가 디스플레이되어 있는데, 본 발명의 특징과 장점이 명백히 도시되어 있다. 먼저, 도 2a를 참조하면, 도 4의 컴퓨터 시스템에서 실행되는 음성 가동 애플리케이션의 대표적인 스크린이 디스플레이되는 스크린(30)이 다시 도시되고 있다. 이 스크린(30)에서, 연관된 툴바 메뉴(35)를 구비하는 통상적인 툴바(34)가 스크린(30)의 왼쪽 위 구석에 나타난다는 것을 알 수 있다. 이 툴바 메뉴(35)에서 나타난 하나의 음성 인식가능 커맨드는 파일 열기(open file)(37)일 수 있다. 단말 사용자가 도 1의 인터페이스의 경우에서와 같이 파일 열기 커맨드(37)를 발음하면, 천숙한 확인 영역(36)이 메뉴(35) 근처에 나타나며, 여기서 음성 인식된 발음인 파일 열기(또는 시스템이 인식한 것이면 무엇이든지)가 나타나게 되어, 사용자의 말이 올바르게 인식되었는지에 관한 피드백을 사용자에게 제공한다.

마찬가지로, 도 2b를 참조하면, 또 다른 일련의 음성 인식가능 커맨드와도 연관된 작업이나 툴바(34)와 연관되어 다른 툴바 메뉴(38)가 스크린(30)의 위부분에 나타날 수 있다. 이 툴바 메뉴는 도 2a의 위치와 다른 스크린(30)상의 위치에 있고, 또한 디스플레이된 인쇄(39)와 같은 다른 음성 가동가능 커맨드를 구비할 수도 있다는 것을 알 수 있다. 도 1 및 도 2a의 경우와 마찬가지로 방식으로, 이 음성 가동가능 커맨드인 인쇄를 사용자가 발음하면, 이 도 2b에서는 확인 영역(36)이 디스플레이 스크린(30) 상의 음성 가동가능 인쇄(39) 커맨드가 나타난 곳 가까이 나타나는 것을 볼 수 있다.

이들 도 2a와 도 2b를 도 1의 종래 기술 사용자 인터페이스와 비교해 보면, 중요하고 의미있는 차이가 드러난다. 전술한 바와 같이 도 1의 종래 기술의 경우, 단말 사용자(32)의 집중 영역은 디폴트 확인 영역(default confirmation area)(34)이 나타난 곳으로부터 공간적으로 상당히 떨어져 있을 수 있으며, 시선이 이 영역(32)과 영역(34) 사이를 가로지르는 동안 시선의 초점을 다시 맞추는 필요가 있게 되는 등, 연관된 바람직하지 못한 집중 혼란을 야기한다.

그러나, 이와는 대조적으로, 도 2a와 도 2b의 사용자 인터페이스의 예에서는 음성 커맨드를 유발하는 (이는 다시 확인 영역(36)이 디스플레이되게 함) 스크린 상의 영역 가까이 있고, 따라서 사용자 시선의 초점이 맞추어져 있을 것으로 생각되는 스크린(30) 상의 영역 내에 이러한 확인 영역(36)이 디스플레이된다는 중요한 차이가 있다는 것을 알게 될 것이다. 따라서, 예컨대 도 2a에서 확인 영역(36)은, 음성 가동가능 파일 열기 커맨드(35)가 툴바 메뉴 내에 디스플레이되는 스크린(30) 상의 위치 가까이, 즉 바로 아래에 있게 되는데, 이 파일 열기 커맨드(35)가 발음되면 스크린 상에 이 확인 영역(36)과 인식된 파일 열기 커맨드가 디스플레이된다. 마찬가지로, 도 2b에서 이 확인 영역(36)은 이제 도 2a의 위치와는 다른 스크린 상의 위치, 즉, 도 2a의 파일 커맨드를 포함하는 툴바 메뉴(35)와는 다른 위치에 나타난 툴바 메뉴(38)와 연관된 음성 가동가능 커맨드인 인쇄(39)와 가까운 위치에 나타난다. 이러한 방식으로, 사용자 시선은 확인 영역(36) 내에 표시된 바와 같이 커맨드가 올바르게 해석되었는지를 확인하기 위해, 그 음성 가동가능 커맨드를 프롬프트한 디스플레이 스크린 상의 위치로부터 혼란스러운 상당한 거리를 가로지르는 필요가 없다.

본 발명의 또 다른 중요한 특징은, 해석된 커맨드를 분석하여 확인 영역(36)에 디스플레이하는 음성 내비

게이션 프로그램 도중에, 필요에 따라 적절한 인공 지능적 요소를 포함할 수 있는 문맥 민감형 분석 기능(a context-sensitive analysis function)이 음성 인식 시스템내에 구축될 수 있어서, 이 확인 영역(36)을 역시 스크린(30)상에 디스플레이되는 커맨드 타겟 영역(이는 확인 영역과, 해석된 커맨드가 디스플레이 되도록 유망한, 즉, 그 원인을 제공한 영역이기도 하다)에 배치할 수 있다는 것이다. 따라서, 예를 들어, 단일 사용자가 파일 열기를 발음하면, 도 4의 시스템 상에서 실행되는 음성 인식 프로그램내의 문맥 민감형 분석 기능 서브루틴(subroutine)이나 프로그램은, 이 파일 열기 커맨드가 스크린(30) 상의 왼쪽 위 구석에 현재 디스플레이되고 있기 때문에 확인 영역(36)은 이 파일 열기 커맨드가 볼바(34)의 드롭다운(drop-down) 메뉴 내에 나타나는 위치와 가까운 곳에 디스플레이되어야 한다고 결정할 것이다. 마찬가지로, 음성 인식 시스템이 인쇄(39) 커맨드가 발음되었음을 감지하고, 이를 이 커맨드가 디스플레이(30)의 오른쪽 위 구석에 디스플레이된다는 사실과 관련시키면, 본 발명의 시스템은 도 2a의 확인 영역(36)의 디스플레이 위치와는 달리, 도 2b의 확인 영역(36)을 스크린 상의 다른 위치 예컨대 단일 사용자로 하여금 인쇄라고 발음하도록 유망한 인쇄(39) 커맨드가 나타난 곳 가까이 있는 위치에 디스플레이하는 것이 이제는 보다 적절하다고 판단할 것이다.

이제, 도 3을 참조하면, 본 발명의 전술한 특징을 구현하기 위해, 도 4의 시스템 상에서 프로그램 코드가 어떻게 실행되는지를 나타내는 흐름도가 도시되어 있다. 우선, 단일 사용자가 도 4의 시스템 상에서 실행될 상자(40)에 도시한 바와 같이, 음성 인식 또는 네비게이션 애플리케이션을 불러낸다고 가정될 것이 다. 도 3의 프로세스를 계속하기 이전에 확인 영역(36)의 지속(persistence)과 소멸(dissipation)을 조절하는 것이 바람직할 것이다. 따라서, 박스(42)에 도시한 바와 같이, 이 프로그램 코드는 이 특징을 수용하도록 적응될 수 있다. 이와는 달리, 본 발명의 특징은 이러한 지속과 소멸을, 예를 들어, 음성 인식 시스템이 단일 사용자의 발음을 인식하는 신뢰도(degree of confidence)의 함수로서 자동으로 선택적으로 변경할 수도 있다는 것을 이해해야 할 것이다. 예컨대, 음성 인식 시스템이 애플리케이션 닫기(close application)라는 단어를 인식하는 신뢰도가 거의 100%이면, 이 피드백은 단일사용자에게 중요하지 않을 것이므로, 확인 메시지의 지속 시간 및 소멸 시간을 실질적으로 줄이는 것이 바람직할 것이다. 한편, 음성 인식 애플리케이션이 그 발음을 인식하는 것이 어렵다면, 확인 메시지의 지속 및 소멸 시간을 자동으로 높이는 것이 바람직할 것이다. 이로 인해, 단일 사용자는 올바르게 해석되지 않은 커맨드를 알아내고, 확인 영역 내에 올바른 원하는 커맨드를 수동으로 입력하거나 키보드 또는 지시 장치를 이용하여 그 커맨드를 수동으로 실행시키거나, 비실행 특징(undo feature)을 불러내는 등의 정정 동작을 취하는데 보다 많은 시간을 쓸 수 있을 것이다.

계속해서 도 3을 참조하면, 일단 음성 인식 애플리케이션 프로그램이 불러져서 실행되면, 박스(44)에 도시한 바와 같이 단일 사용자가 디스플레이(30) 상에 디스플레이된 상관 커맨드(correlative command)를 보고 이에 응답하여 음성 인식 커맨드를 입력한다고 가정될 것이다. 이러한 입력은 도 4의 모니터 상에서 발생하는 하나 이상의 디스플레이 스크린(30)을 검토하고 이에 응답하여 도 4의 시스템에서 도시한 마이크로폰(28)을 이용하여 이루어질 수 있다.

일단 이 음성 인식 커맨드가 발음되면, 도 4의 시스템은 이어서, 확인 영역(36) 내에서 단일 사용자에게 피드백을 제공하기 위한 적절한 공간적 위치를 검출할 것인데, 이 단계는 도 3의 블록(46)으로 나타내었다. 이 기능 블록의 구현은 전술한 바와 같이 사용자가 이 애플리케이션과 이전에는 어떻게 상호작용했는지를 감지 내지 참조하는 인공 지능 요소 등 적절한 다양한 요소를 포함할 수 있다는 것을 상기하여야 한다. 또한, 블록(46)에서의 이러한 검출은 어떤 음성 가능한 커맨드가 디스플레이 스크린(30) 상에서 현재 디스플레이되고 있는가를 더 분석하는 것을 포함할 수도 있고, 음성 가능한 프로그램과의 이전 상호작용의 분석에 기초하여 디스플레이 스크린(30) 상에서 단일 사용자의 시선과 지능적 초점이 다음에 어디에서 나타날 것인가에 관한 지능적 예측을 포함할 수도 있다.

계속해서 도 3을 참조하면, 일단 프로그램 코드가 확인 영역(36) 내의 피드백을 위한 바람직한 공간적 위치를 검출하고 나면, 이어서 시스템은 박스(46)의 기능을 참조해서 이전에 이루어진 분석에 기초하여 (도 2a 내지 도 2b의) 스크린(30) 상에 위치한 확인 영역(36) 내에 적절한 피드백 메시지를 디스플레이할 것이다(48). 적절한 피드백 메시지의 하나는 단일 사용자로부터의 구술 커맨드나 인스트럭션을 음성 인식 시스템에 의해 해석된 대로 문자와 숫자로 단순히 디스플레이하는 것임을 알 수 있을 것이다. 예컨대, 단일 사용자가 닫기 메뉴(35) 내에서 음성 가능한 커맨드로서 파일 열기가 있음을 확인하고 이에 응답하여 파일 열기라는 단어를 발음했다고 판단되는 경우, 음성 인식 시스템은 확인 영역(36) 내에 파일 열기라는 단어를 그대로 디스플레이할 수 있다.

발음된 음성 커맨드의 최선의 추측 해석이 디스플레이된 후(48), 도 3의 프로세스는 도 3의 결정 블록(50)에 도시한 바와 같이 다른 메시지나 커맨드가 단일 사용자에게 의해 발음되었는지를 문의한다. 다른 메시지가 검출되었으면, 흐름은 블록(50)의 왼쪽으로 나가서 경로(52)를 따라 다음 음성 커맨드 발음 입력이 분석될 블록(44)으로 돌아간다. 이어서, 흐름은 도 3의 흐름도를 따라 아래로 계속되는데, 여기에서는 다음 발음의 확인 영역(36)을 위한 다음 원하는 공간적 위치가 검출되고 특정 커맨드가 분석된다. 한편, 결정 블록(50)의 문의에 응답하여, (예를 들어, 음성 네비게이션 애플리케이션이 종료하여) 다른 음성 커맨드가 입력되지 않았거나 입력되지 않을 것이라고 판단되면, 프로세스는 결정 블록(50)의 오른쪽으로 나가서 프로세스가 종료(54)한다.

본 발명이 유리하게 채용되는 컴퓨터 시스템의 블록도를 설명하기 이전에, 몇 가지 사실에 주목해야 한다. 확인 영역(36)이 적절한 지속 및 소멸 길이 이후에 사라져도, 커맨드가 올바르게 인식되면 그 커맨드는 시스템에 의해 자동으로 실행될 것이라는 사실은 전술한 바로부터 명백하다. 한편, 커맨드가 올바르게 인식되지 못하면, 확인 영역이 지속되어 소멸되기 이전이라도, 단일 사용자는 (사용자가 결정된 선호도(user-determined preference)에 따라, 키보드로의 수동 입력, 필요하다면 지시 장치를 이용하거나 중지(STOP)와 같은 특정 음성 커맨드를 발음함으로써 인식된 커맨드를 무효로 할 수 있다.

전술한 설명에서, 음성 인식은 단일 사용자가 디스플레이된 볼바 등으로부터 선택하는 사전결정된 커맨드 선택에 대해 채용되었지만, 본 발명을 이에 제한시키려는 의도가 아니라는 것을 추가로 지적하고자 한다. 따라서, 스크린 상에 나타난 메시지와 연관되지 않은 음성 가능한 동작이 존재할 수도 있다. 이러한 대표적인 예로는 커서를 아래로 움직이는 것과 같은 방향 제어 또는 그림을 확대하는 것이 있을

수 있다. 이렇게 발음된 방향도 스크린 상에 디스플레이된 커맨드 선택 메뉴와 유사한 방식으로 음성 시스템에 의해 인식될 수 있고 음성 시스템이 이에 따라 작동할 수 있으며, 이 때 인식된 커맨드를 실행하거나 이를 키보드 또는 지시 장치를 이용하여 정정하기 이전에 발음된 단어의 확인이 확인 영역(36) 내에 표시될 수 있다. 또한, 음성 커맨드를 선택할 메뉴가 존재하지 않는 경우에도 본 발명의 근본적인 개념을 유지하면서, 음성 인식 시스템이 단일 사용자에게 의해 발음된 단어에 의존하여 확인 영역(36)을 배치할 장소를 지능적으로 결정할 수 있다는 점에서, 본 음성 인식 시스템은 커맨드-문맥 민감형(command-context sensitive)이라고 간주된다. 이를테면, 단일 사용자가 커서를 오른쪽으로 2 인치만큼 이동시키라고 발음한 경우, 시스템은 커서라는 단어를 인식함으로써 커서와 연관된 이동, 형태 변화 등과 같은 어떤 동작이 요구된다는 것을 판단할 수 있다. 따라서, 이 시스템은 디스플레이 스크린(30) 상의 커서의 현재 위치 가까이 디스플레이된 인식된 커맨드와 함께 확인 영역(36)을 자동으로 디스플레이할 것이다. 물론, 이는 단일 사용자가 자신의 시선을 커서 상의 커맨드 바로 앞에 초점을 맞출 것이고 따라서 확인 메시지가 현재 커서 위치 근처 머당가에 나타나리라고 기대하는 가능성을 인식함에 따른 것이다.

도 4는 본 발명의 개선된 지시 장치 분석 시스템(pointing device resolution system)을 유리하게 채용하는 컴퓨터 시스템의 바람직한 실시예를 도시하고 있다. 이 시스템은 CPU(10), 롬(ROM : Read Only Memory)(11), 램(Random Access Memory : RAM)(12), I/O 어댑터(13), 사용자 인터페이스 어댑터(18), 통신 어댑터(14), 디스플레이 어댑터(19)를 포함하고 있는데, 이들 모두는 공통 어드레스/데이터 및 제어 버스 또는 버스(16)를 통해 상호접속되어 있다. 이 구성요소 각각은 시스템 내의 각 구성요소에 특정 어드레스 범위를 제공하는 것과 같은 방법을 포함하는 단일자에게 알려진 통상적인 기법을 이용하여 공통 버스를 액세스하는데, 이때 버스마스터(busmaster)의 역할은 CPU가 한다. 도 4에서 더 볼 수 있듯이, DASD(15)와 같은 외부 장치는 I/O 어댑터(13)와 같은 각각의 어댑터를 통해 공통 버스(16)로 인터페이스된다. 디스플레이(21)와 같은 다른 외부 장치는 디스플레이 어댑터(19)와 같은 그들 각각의 어댑터를 이용하여 버스(16)와 디스플레이(21) 또는 다른 장치 사이의 데이터 흐름을 제공한다. 다양한 사용자 인터페이스 수단이 상호접속을 위해 또한 사용자 인터페이스 어댑터(18)와 더불어 사용하기 위해 제공된다. 도면에서, 사용자 인터페이스 어댑터(18)에는 조이스틱(23), 마우스(25), 키보드(17), 스피커 및/또는 마이크로폰(27)과 같은 대표적인 사용자 입력 디바이스가 부착되어 있다. 이들 각각의 유닛은 주지의 것이므로 본 명세서에서는 설명하지 않는다.

본 발명은 본질적으로 RS/6000, RISC 기반 워크스테이션(workstations), AIX와 OS/2 운영 체제를 각각 실행하는 IBM사의 개인용 컴퓨터, 또는 다른 회사의 유사한 마신과 같은 임의의 컴퓨터 시스템과 대응 마이크로프로세서 상에서 구현될 수 있는데, 이들은 예를 들어, RS/6000 워크스테이션의 경우에는 604 PowerPC RISC 칩을 포함한다. (RS/6000, IBM, AIX, OS/2, PowerPC는 IBM사의 등록 상표이다.)

도 4의 CPU(10) 내에는 도 4의 시스템의 올바른 동작에 필요한 시스템 어드레스, 데이터, 제어 처리 기능을 수행하는 하나 이상의 마이크로프로세서가 포함되어 있다. 본 발명이 다양한 마이크로프로세서 설계로의 응용을 허용하지만, 본 명세서에 개시된 실시예에서 마이크로프로세서는 IBM 코퍼레이션에서 제조한 PowerPC 604 마이크로프로세서 형태를 취하는데, 이는 감축형 인스트럭션 세트 컴퓨터(Reduced Instruction Set Computer : RISC) 마이크로프로세서로 알려져 있는 마이크로프로세서의 일종이다. 이러한 마이크로프로세서의 구조와 동작에 관한 세부 사항은 PowerPC 604 RISC Microprocessor Users Manual, Document # MPC604UM/AD, 1994년 12월(저작권자는 IBM사임)로부터 얻을 수 있는데, 이를 본 명세서에서 참조로서 인용한다.

본 발명의 맥락에서, 사용자는 마우스(25)와 음성 자동 네비게이션과 같은 다양한 지시 장치를 이용하여 조작할 수 있는 디스플레이(21) 상의 커서와 합입 또는 팝다운 메뉴와 같은 다양한 객체를 관찰할 수 있다. 사용자 인터페이스 어댑터(18)와 연관된 프로그램 코드는 램(12) 및/또는 DASD(15) 내에 상주하는 오퍼레이팅 환경 및 애플리케이션 코드와 함께 지시 장치(25)와 마이크로폰(27)을 위한 디바이스 드라이버를 경우하여, 마이크로폰(27)으로 입력된 상판 음성 커맨드에 따라 또한 이와 관련하여 디스플레이(21) 상의 커서의 움직임을 촉진시키거나 가능하게 할 것이다.

전술한 상세한 설명으로부터, 본 발명의 진정한 사상을 벗어나지 않는 범위 내에서 본 발명의 바람직한 실시예의 다양한 수정과 변경이 가능하다는 것을 이해할 것이다. 이 명세서는 설명을 위한 것이며, 제한하려는 의도로 해석되어서는 안 된다. 본 발명의 범주는 후속하는 청구 범위에 의해서만 제한되어야 한다.

발명의 효과

본 발명에 의하면 디스플레이 스크린 상의 관심 내용을 가리거나 및/또는 디스플레이 스크린 상의 현재 관심 위치로부터 확인 영역이 디스플레이되는 다른 위치로 시선을 이동시키도록 요구함으로써 단일 사용자 초점을 훼손하는 고정 위치 확인 영역과 연관된 산만성(distractions)이 방지된다.

(5) 청구의 범위

청구항 1

음성 인식 애플리케이션(a voice recognition application)을 실행하는 음성 인식 컴퓨터 시스템에서 디스플레이 상에 피드백을 제시하는 방법에 있어서,

① 음성 커맨드를 감지하는 단계와,

② 상기 감지된 음성 커맨드의 함수로서 상기 디스플레이 상의 다수의 위치(positions) 중 하나의 위치를 선택하는 단계와,

③ 상기 다수의 위치 중 상기 하나의 위치에 가까이 있고 상기 커맨드에 대응하는 확인 영역(a

confirmation area)에 피드백을 디스플레이하는 단계
를 포함하는 디스플레이 상의 피드백 제시 방법.

청구항 2

제 1 항에 있어서,

④ 상기 디스플레이 이후의 시간에 상기 피드백을 소멸(extinguish)시키는 단계를 더 포함하는 디스플레이 상의 피드백 제시 방법.

청구항 3

제 2 항에 있어서,

⑤ 상기 디스플레이 이후에 상기 커맨드를 자동으로 실행하는 단계를 더 포함하는 디스플레이 상의 피드백 제시 방법.

청구항 4

제 3 항에 있어서,

상기 소멸 시간은 동적으로 선택되고, 상기 피드백은 상기 디스플레이 상의 색채 변화, 애니메이션(animation), 메시지, 또는 영상 속소를 포함하는 그룹으로부터 선택되는 디스플레이 상의 피드백 제시 방법.

청구항 5

제 3 항에 있어서,

⑥ 상기 감지의 정확도에 대응하는 신뢰 요소(a confidence factor)를 결정하는 단계 - 상기 소멸 시간은 상기 신뢰 요소에 대응함 - 를 더 포함하는 디스플레이 상의 피드백 제시 방법.

청구항 6

제 5 항에 있어서,

⑦ 상기 디스플레이 상의 장소(a location)에서 상기 음성 커맨드에 대응하는 상기 디스플레이 정보를 표시하는 단계 - 상기 음성 커맨드는 상기 정보에 대응하고, 상기 다수의 위치 중 상기 하나의 위치는 상기 장소에 대응함 - 를 더 포함하는 디스플레이 상의 피드백 제시 방법.

청구항 7

제 5 항에 있어서,

⑧ 상기 음성 커맨드를 분석하는 단계 - 상기 다수의 위치 중 상기 하나의 위치는 상기 분석의 함수임 - 를 더 포함하는 디스플레이 상의 피드백 제시 방법.

청구항 8

제 5 항에 있어서,

상기 컴퓨터 시스템은 상기 디스플레이 상에 그래픽 사용자 인터페이스(a graphical user interface)와, 지시 장치(a pointing device)를 포함하고, 상기 음성 커맨드는 상기 지시 장치에 의해 구동가능한 상기 그래픽 사용자 인터페이스 상에 디스플레이된 커맨드에 대응하는 디스플레이 상의 피드백 제시 방법.

청구항 9

제 1 항에 있어서,

⑨ 다음 음성 커맨드를 감지하는 단계와,
⑩ 상기 감지된 다음음성 커맨드의 함수로서 상기 디스플레이 스크린 상의 다수의 위치 중 다음 위치를 선택하는 단계와,
⑪ 상기 다수의 위치 중 상기 다음 위치 가까이 있고 상기 다음 음성 커맨드에 대응하는 다음 확인 영역에 다음 피드백을 디스플레이하는 단계
를 더 포함하는 디스플레이 상의 피드백 제시 방법.

청구항 10

제 5 항에 있어서,
상기 신뢰 요소와 상기 소멸 시간은 반비례 관계(inversely related)인 디스플레이 상의 피드백 제시 방법.

청구항 11

음성 인식 애플리케이션을 실행하는 음성 인식 컴퓨터 시스템에서 디스플레이 상에 피드백을 제시하는 장치에 있어서,

- ① 음성 커맨드를 감지하는 수단과,
 - ② 상기 감지된 음성 커맨드의 함수로서 상기 디스플레이 상의 다수의 위치 중 하나의 위치를 선택하는 수단과,
 - ③ 상기 다수의 위치 중 상기 하나의 위치에 가까이 있고 상기 커맨드에 대응하는 확인 영역에 피드백을 디스플레이하는 수단
- 을 포함하는 디스플레이 상의 피드백 제시 장치.

청구항 12

제 11 항에 있어서,

- ④ 상기 디스플레이 이후의 시간에 상기 피드백을 소멸시키는 수단을 더 포함하는 디스플레이 상의 피드백 제시 장치.

청구항 13

제 12 항에 있어서,

- ⑤ 상기 디스플레이 이후에 상기 커맨드를 자동으로 실행하는 수단을 더 포함하는 디스플레이 상의 피드백 제시 장치.

청구항 14

제 13 항에 있어서,

상기 소멸 시간은 동적으로 선택되고 상기 피드백은 상기 디스플레이 상의 색채 변화, 애니메이션, 메시지, 또는 영상 축소를 포함하는 그룹으로부터 선택되는 디스플레이 상의 피드백 제시 장치.

청구항 15

제 13 항에 있어서,

- ⑥ 상기 감지의 정확도에 대응하는 신뢰 요소를 결정하는 수단 - 상기 소멸 시간은 상기 신뢰 요소에 대응함 - 을 더 포함하는 디스플레이 상의 피드백 제시 장치.

청구항 16

제 15 항에 있어서,

- ⑦ 상기 디스플레이 상의 장소에서 상기 음성 커맨드에 대응하는 상기 디스플레이 정보를 묘사하는 수단 - 상기 음성 커맨드는 상기 정보에 대응하고, 상기 다수의 위치 중 상기 하나의 위치는 상기 장소에 대응함 - 을 더 포함하는 디스플레이 상의 피드백 제시 장치.

청구항 17

제 15 항에 있어서,

- ⑧ 상기 음성 커맨드를 분석하는 수단 - 상기 위치 중 상기 하나는 상기 분석의 함수임 - 을 더 포함하는 디스플레이 상의 피드백 제시 장치.

청구항 18

제 15 항에 있어서,

상기 컴퓨터 시스템은 상기 디스플레이 상에 그래픽 사용자 인터페이스와 지시 장치를 포함하고, 상기 음성 커맨드는 상기 지시 장치에 의해 구동가능한 상기 그래픽 사용자 인터페이스 상에 디스플레이된 커맨드에 대응하는 디스플레이 상의 피드백 제시 장치.

청구항 19

제 15 항에 있어서,

상기 컴퓨터 시스템은 상기 디스플레이 상에 그래픽 사용자 인터페이스와 지시 장치를 포함하고, 상기 음성 커맨드는 상기 그래픽 환경 상에서 실행되는 애플리케이션에 의해 해석된 비가시(non-visible) 커맨드에 대응하는 디스플레이 상의 피드백 제시 장치.

청구항 20

제 11 항에 있어서,

- ㉠ 다음 음성 커맨드를 감지하는 수단과,
 - ㉡ 상기 감지된 다음 음성 커맨드의 함수로서 상기 디스플레이 스크린 상의 다수의 위치 중 다음 위치를 선택하는 수단과,
 - ㉢ 상기 다수의 위치 중 상기 다음 위치 가까이 있고 상기 다음 음성 커맨드에 대응하는 다음 확인 영역에 다음 피드백을 디스플레이하는 수단
- 을 더 포함하는 디스플레이 상의 피드백 제시 장치.

청구항 21

제 15 항에 있어서,

상기 신뢰 요소와 상기 소멸 시간은 비례 관계에 있고, 상기 컴퓨터 시스템은 상기 디스플레이 상에 그래픽 사용자 인터페이스와 지시 장치를 포함하며, 상기 음성 커맨드는 상기 그래픽 사용자 인터페이스에서 실행되는 애플리케이션에 의해 해석된 비가시 커맨드에 대응하는 디스플레이 상의 피드백 제시 장치.

청구항 22

음성 인식 애플리케이션을 실행하는 음성 인식 컴퓨터 시스템에서 디스플레이 상에 피드백을 제시하는 프로그램 제품에 있어서,

- ㉠ 음성 커맨드를 감지하는 프로그램 코드 수단과,
 - ㉡ 상기 감지된 음성 커맨드의 함수로서 상기 디스플레이 상의 다수의 위치 중 하나의 위치를 선택하는 프로그램 코드 수단과,
 - ㉢ 상기 다수의 위치 중 상기 하나의 위치에 가까이 있고 상기 커맨드에 대응하는 확인 영역에 피드백을 디스플레이하는 프로그램 코드 수단
- 을 포함하는 디스플레이 상의 피드백 제시 프로그램 제품.

청구항 23

제 22 항에 있어서,

- ㉣ 상기 디스플레이 이후의 시간에 상기 피드백을 소멸시키는 프로그램 코드 수단을 더 포함하는 디스플레이 상의 피드백 제시 프로그램 제품.

청구항 24

제 23 항에 있어서,

- ㉤ 상기 디스플레이 이후에 상기 커맨드를 자동으로 실행하는 프로그램 코드 수단을 더 포함하는 디스플레이 상의 피드백 제시 프로그램 코드 제품.

청구항 25

제 24 항에 있어서,

상기 소멸 시간은 동적으로 선택되고 상기 피드백은 상기 디스플레이 상의 색채 변화, 애니메이션, 메시

지, 또는 영상 축소를 포함하는 그룹으로부터 선택되는 디스플레이 상의 피드백 제시 프로그램 제품.

청구항 26

제 24 항에 있어서,

⑥ 상기 감지의 정확도에 대응하는 신뢰 요소(a confidence factor)를 결정하는 프로그램 코드 수단 - 상기 소멸 시간은 상기 신뢰 요소에 대응함 - 을 더 포함하는 디스플레이 상의 피드백 제시 프로그램 제품.

청구항 27

제 26 항에 있어서,

⑦ 상기 디스플레이 상의 장소에서 상기 음성 커맨드에 대응하는 상기 디스플레이 정보를 묘사하는 프로그램 코드 수단 - 상기 음성 커맨드는 상기 정보에 대응하고, 상기 다수의 위치 중 상기 하나의 위치는 상기 장소에 대응함 - 을 더 포함하는 디스플레이 상의 피드백 제시 프로그램 제품.

청구항 28

제 26 항에 있어서,

⑧ 상기 음성 커맨드를 분석하는 프로그램 코드 수단 - 상기 다수의 위치 중 상기 하나의 위치는 상기 분석의 함수임 - 을 더 포함하는 디스플레이 상의 피드백 제시 프로그램 제품.

청구항 29

제 26 항에 있어서,

상기 컴퓨터 시스템은 상기 디스플레이 상에 그래픽 사용자 인터페이스와 지시 장치를 포함하고, 상기 음성 커맨드는 상기 지시 장치에 의해 구동가능한 상기 그래픽 사용자 인터페이스 상에 디스플레이된 커맨드에 대응하는 디스플레이 상의 피드백 제시 프로그램 제품.

청구항 30

제 26 항에 있어서,

상기 컴퓨터 시스템은 상기 디스플레이 상에 그래픽 사용자 인터페이스와 지시 장치를 포함하고, 상기 음성 커맨드는 상기 그래픽 환경 상에서 실행되는 애플리케이션에 의해 해석된 비가시 커맨드에 대응하는 디스플레이 상의 피드백 제시 프로그램 제품.

청구항 31

제 22 항에 있어서,

⑨ 다음 음성 커맨드를 감지하는 프로그램 코드 수단과,

⑩ 상기 감지된 다음 음성 커맨드의 함수로서 상기 디스플레이 스크린 상의 다수의 위치 중 다음 위치를 선택하는 프로그램 코드 수단과,

⑪ 상기 다수의 위치 중 상기 다음 위치 가까이 있고 상기 다음 음성 커맨드에 대응하는 다음 확인 영역에 다음 피드백을 디스플레이하는 프로그램 코드 수단

을 더 포함하는 디스플레이 상의 피드백 제시 프로그램 제품.

청구항 32

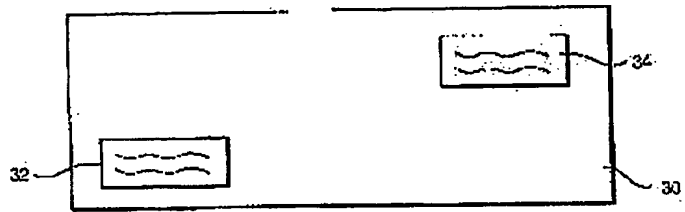
제 26 항에 있어서,

상기 신뢰 요소와 상기 소멸 시간은 반비례 관계인 디스플레이 상의 피드백 제시 프로그램 제품.

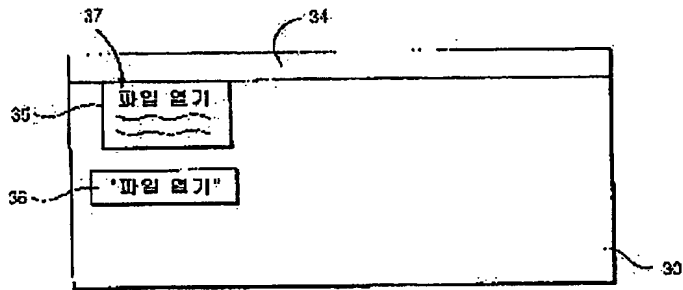
도면

도면1

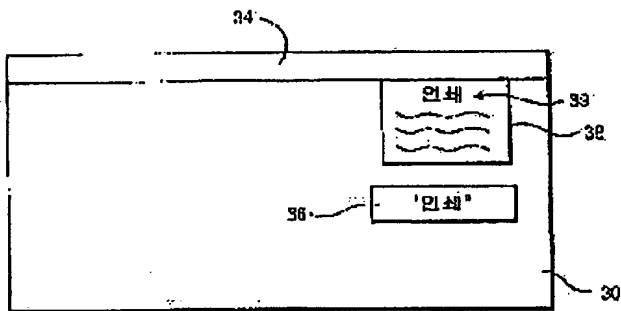
(종래기술)



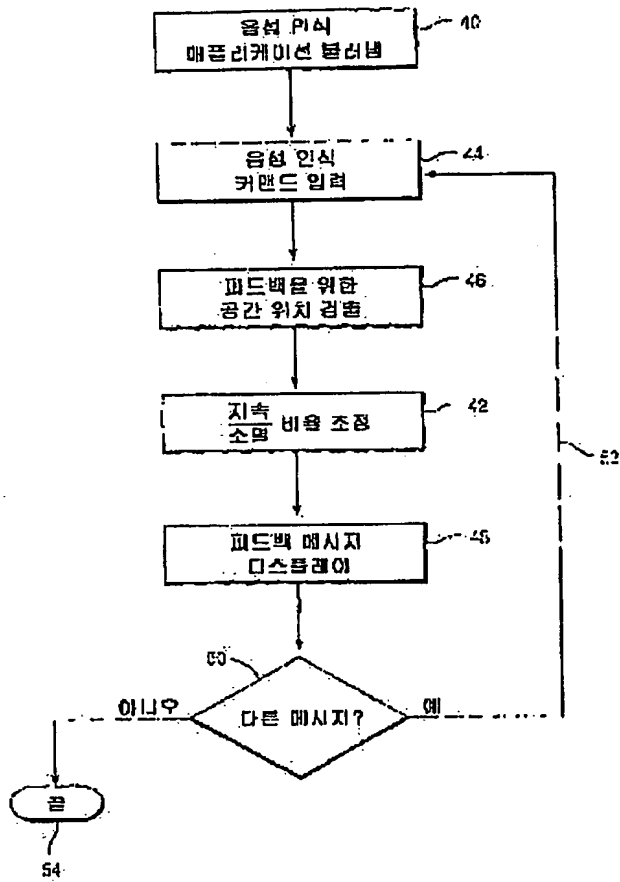
도면2a



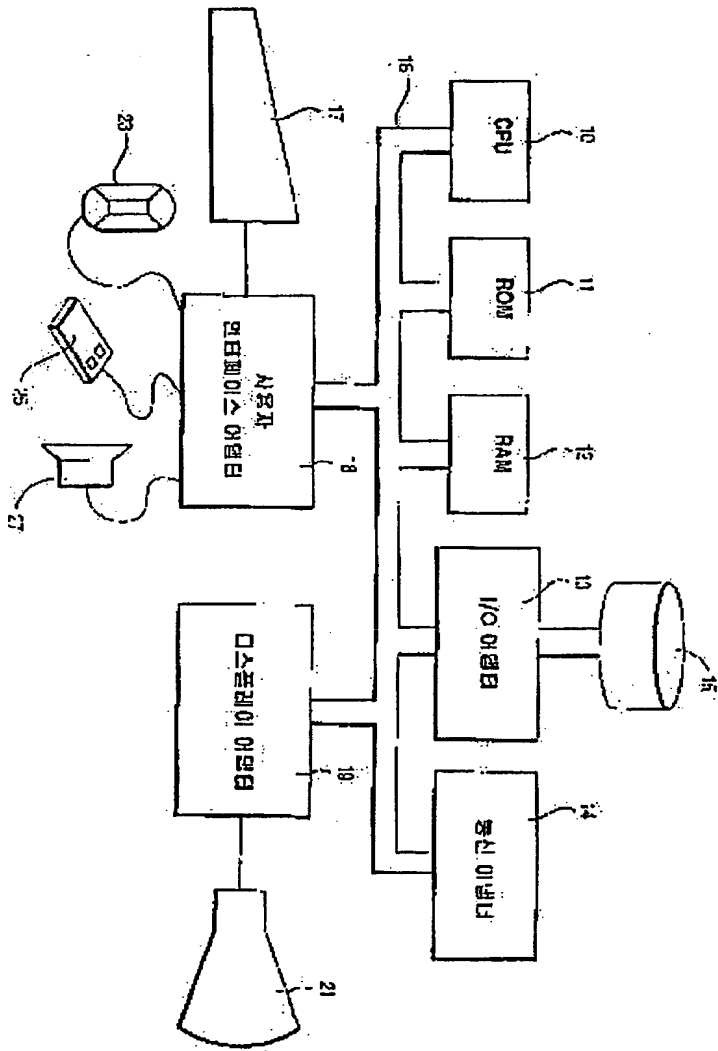
도면2b



도면3



도면4



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.